

Towards OpenMath Content Dictionaries as Linked Data

Christoph Lange

Computer Science, Jacobs University Bremen,
ch.lange@jacobs-university.de

Abstract. “The term ‘Linked Data’ refers to a set of best practices for publishing and connecting structured data on the web” [7]. Linked Data make the Semantic Web work practically, which means that information can be retrieved without complicated lookup mechanisms, that a lightweight semantics enables scalable reasoning, and that the decentral nature of the Web is respected. OpenMath Content Dictionaries (CDs) have the same characteristics – *in principle*, but not yet in practice. The Linking Open Data movement has made a considerable practical impact: Governments, broadcasting stations, scientific publishers, and many more actors are already contributing to the “Web of Data”. Queries can be answered in a distributed way, and services aggregating data from different sources are replacing hard-coded mashups. However, these services are currently entirely lacking mathematical functionality. I will discuss real-world scenarios, where today’s RDF-based Linked Data do not quite get their job done, but where an integration of OpenMath *would* help – were it not for certain conceptual and practical restrictions. I will point out conceptual shortcomings in the OpenMath 2 specification and common bad practices in publishing CDs and then propose concrete steps to overcome them and to contribute OpenMath CDs to the Web of Data.

1 Linked Data State of the Art

The Linked Data principles, established by BERNERS-LEE in 2006 [4] consist of four simple rules for publishing machine-understandable data on the web¹:

1. Use URIs to identify things.
2. Use HTTP URIs so that these things can be referred to and looked up (“dereferenced”) by people and user agents.²
3. Provide useful³ information about the thing when its URI is dereferenced, using standard formats such as RDF/XML.
4. Include links to other, related URIs in the exposed data to improve discovery of other related information on the Web.

¹ here cited as paraphrased by Wikipedia [29]

² I. e., the URI is treated as a URL.

³ This usually means: machine-understandable.

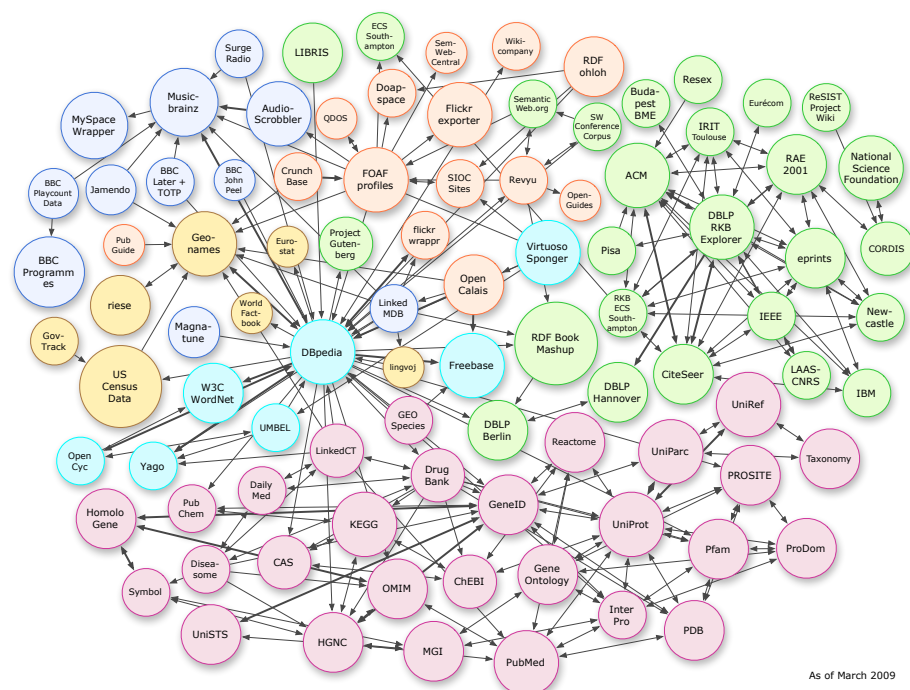


Fig. 1. Linked Open Datasets as of March 2009 [11]

These principles are widely considered to have made the Semantic Web vision work practically. A lot of providers have already published their data according to these principles and interlinked them with other datasets (cf. figure 1). The hub in this big picture is DBpedia, a huge collection of general-purpose data extracted from Wikipedia and made available as RDF. Data from specific domains, such as scientific publications (green), biomedicine (pink), social networks (orange), multimedia (dark blue), and government statistics (yellow) have also been published as Linked Open Data. Linked Data do not *have* to be open⁴, but making datasets open of course helps to interlink and reuse knowledge; therefore, the open datasets have so far been the most visible and most widely used instance of Linked Data. Applications include browsers, which allow users to traverse the Web of Data and discover connections, semantic search engines and indexes, which enable a more accurate information retrieval than keyword-based engines, as well as mashups that aggregate Linked Data from distributed sources and expose them via a coherent user interface (see, e. g., [15] for an interactive map of database researchers and their publications, filterable by research topics).

⁴ In fact they can also be useful in intranet settings, cf. [24]

Listing 1.1. Geese on the Isle of Wight, RDF data in Turtle notation, from data.gov.uk (all URIs abbreviated, namespace prefix mappings omitted for brevity; see [28] for full example)

```

ahs:EH100                                # just some ID for this data point
  scv:dimension env:isle-of-wight ;      # the "region" dimension
  scv:dimension env:year-2008 ;          # the "time" dimension
  scv:dimension env:geese ;              # the type of items counted
  rdf:value      "693"^^xsd:decimal ;   # the count
  scv:dataset    ahs2:livestock .        # back-reference to the dataset

```

2 The Need for Mathematical Semantics

None of the Linked Open datasets and applications known to date deals with mathematical knowledge, not counting mere descriptions that do not involve any mathematical semantics, e. g., of mathematical publications or mathematical research topics, as they can be found in publication datasets or DBpedia. With “Linked Open Numbers” [27], one mathematical dataset has been published, but that was not to be taken serious. There, every natural number from 1 to 999,999,999 is described with its predecessor, successor, natural logarithm, and its name in various natural languages. This is pretty useless information⁵, and indeed “Linked Open Numbers” was an April fool’s joke cartooning the rampant bad habit of mindlessly publishing datasets that are very large but not reasonable at all.

There is, however, no doubt, that mathematical semantics is needed in order to improve, or even enable, certain *serious* applications of Linked Data. I consider statistical datasets, which are now being published as RDF Linked Data, e. g., by the UK and US governments, a prime example. OMITOLA et al. have, for example, used such data in order to answer queries for public sector information in the user’s home region by aggregating data about, e. g., political representatives of the local constituencies, crime statistics for the local county, and waiting list statistics of local hospitals [21]. At the moment, these datasets contain a lot of data points (e. g. the number of geese on the Isle of Wight in 2008; cf. listing 1.1), without making their origin semantically explicit. We have proposed an extension of the relevant Statistical Core Vocabulary (SCOVO), which allows to express the latter knowledge, saying, e. g., “the things that we are counting here are geese (e. g. by referencing <http://dbpedia.org/resource/Goose>) per area and per year” [28]. Mathematical knowledge becomes relevant when modeling *derived values*, such as the geese population density of a region in a given year, defined as the number of geese divided by area.⁶ At the moment, there

⁵ On the other hand, it *might* be useful to publish as Linked Data facts about numbers that are *hard* to compute, e. g. factorizations of large numbers.

⁶ The geese population density is a fictitious example, but in the actual datasets, there are derived values such as the [human] population density of various census regions, or the average number of jobs per citizen.

Listing 1.2. Geese population density of the Isle of Wight, with its mathematical semantics

```
# the density is computed by ...
ahs:PD100 sl:computedFrom [
  # ... calling OpenMath's arith1#divide
  sl:function <http://www.openmath.org/cd/arith1#divide> ;
  sl:arguments
    # ... passing the value of the EH100 data point as first argument
    [ sl:argPosition "1"^^xsd:int ;
      sl:argValue ahs:EH100 ] ,
    # ... and the value of the AR100 data point as second argument
    [ sl:argPosition "2"^^xsd:int ;
      sl:argValue ahs:AR100 ] ] .
```

are a lot of derived values in the datasets published, simply given as additional raw data points. For a client consuming these data, there is no way of verifying their correctness or applying the same derivation rule to new or changed base values, because the derivation rule is not made explicit. We have shown how to make their mathematical semantics explicit – first on the instance level, as that integrates most easily into existing datasets. Let the data point with the ID *ahs:AR100* be the area of the Isle of Wight, and let *ahs:PD100* be the geese population density of the Isle of Wight in 2008, then we could express the fact that the latter is *ahs:EH100* divided by *ahs:PD100* by referencing the OpenMath symbol for division (cf. listing 1.2). In a second step, the same could be done on vocabulary level: In addition to, or alternatively to, explicitly representing the derivation of each data point, one could model a general rule that “for each data point *p* containing a ‘population’ of some region *r* at some point *t* in time and for each data point *a* containing the area of *r* [at time *t*], the population density *d* of *r* at time *t* is defined as $d := \frac{p}{a}$ ”. Recall, however, that the semantics of Linked Data vocabularies is usually intentionally weak in order to enable large-scale applications. Such general rules would require more powerful clients and query engines and might therefore not work as universally as semantically more lightweight (albeit blatantly redundant) annotations of individual data points.

For computing such a derivation, a Linked Data client has to translate these RDF data to an OpenMath object, which has to be fed to a computation service, e.g. a service that speaks SCSCP [12, 13]. We have detailed the translation in [28]. For standard symbols, such as *arith1#divide* here, the translation is pretty straightforward. Computing the division should not be a problem for any OpenMath-aware service, as there is certainly a phrasebook mapping *arith1#divide* to the native division operator of some computer algebra system.

But now suppose that there are more complex, non-standard derivations in our statistical dataset. This makes the case for publishing OpenMath CDs as Linked Data, by the following considerations: Suppose the dataset contains the

Human Development Index (HDI) of a country⁷. Assuming that the four required auxiliary data points have already been computed (LE = life expectancy index, ALI = adult literacy index, GEI = gross enrollment index, and GDP = an index computed from the gross domestic product per capita at purchasing power parity, all normalized to a scale between 0 and 1), the HDI is defined as $\frac{1}{3}(LE + \frac{2}{3}ALI + \frac{1}{3}GEI + GDP)$. In [28], we propose that the dataset publishers define the HDI and its derivation as a symbol in an OpenMath CD that accompanies the dataset, e. g. <http://example.org/statistics>. Now suppose there is a derived data point annotated as `sl:computedFrom [sl:function <http://example.org/statistics#hdi> ; ...]` in analogy to listing 1.2. As OpenMath-based computation services and thus phrasebooks are developed independently from datasets being published, we have little to no chance to expect a phrasebook supporting the <http://example.org/statistics> CD. Therefore, we propose to add support for processing OpenMath CDs to Linked Data clients. For (re)computing an HDI data point derived from four other data points containing the LE , ALI , GEI , and GDP values, the client would download the definition of the <http://example.org/statistics#hdi> symbol from the CD, expand the mathematical expression using the definition, and then send that expanded expression, which only uses operators from the universally understood *arith1* CD, to the computation service.⁸

So far, I have outlined one use case, where OpenMath CDs as Linked Data would be needed. In the following section, I will point out what actions on the OpenMath side that requires. Note that, while the Linked Data principles have been devised in the context of RDF, and while all contemporary Linked Open datasets are available as RDF, the Linked Data guidelines do not prescribe RDF. In fact, RDF might not be the most appropriate representation for mathematical objects. It is at least quite cumbersome to break the ordered tree structure of mathematical expressions down to unordered RDF triples (cf. [19] for one never-adopted suggestion on how that could be done, and [28] for a critical review). For the remainder of this paper, I assume that CDs will be published in their reference XML encoding.

3 Linked Data Principles in OpenMath

First, let us see how much the Linked Data principles cited in section 1 are already respected in the practice of publishing OpenMath CDs:

1. Hardly any CD author uses *CDBase*, which indicates a lack of awareness that things *can* be identified by URIs.
2. The URIs used for OpenMath CDs/symbols are always HTTP URLs, but due to the inconsequent usage of *CDBase* (cf. principle 1), most published

⁷ http://en.wikipedia.org/wiki/Human_Development_Index

⁸ Here, we assume that those values, from which the HDI is computed, are either hard-coded in the dataset, or that they have been computed before, using the same method.

CDs have the base URI <http://www.openmath.org/cd> (i.e. the default value for *CDBase*). Even when disregarding the following principles, this is at least bad style, as the authors who publish such CDs usually do not have control over the [openmath.org](http://www.openmath.org) domain.

3. If people are aware of the fact that OpenMath CDs and symbols have a URI, they usually merely consider it a globally unique *name*, but not a means of locating information about a CD or symbol. At [openmath.org](http://www.openmath.org), the URIs/URLs of CDs at least redirect to human-readable HTML renderings (e.g. <http://www.openmath.org/cd/arith1#plus>~<http://www.openmath.org/cd/arith1.xhtml#plus>), but that does not help a *machine* that is interested in a description of a symbol: Neither is the HTML semantically annotated (e.g. with RDFa, or with parallel markup in the case of rendered MathML formulæ), nor is the CD in its original XML encoding available from that URI in a way that would not require human brain-power.
4. OpenMath CDs are not integrated into the Web of Data at all. Some of the standard CDs make references to mathematical literature, e.g. sections of the Handbook of Mathematical Functions by ABRAMOWITZ and STEGUN [1]. Hyperlinks to its digital counterpart, the Digital Library of Mathematical Functions (DLMF [20]), would be more appropriate here.⁹ Other than that, I am not aware of other links in CDs. However, links to background information about certain mathematical operators or functions, e.g. to the DBpedia editions of their Wikipedia articles, would make sense. Conversely, backlinks from DBpedia to OpenMath CDs would make sense (“go there if you want a description of the mathematical semantics of <http://dbpedia.org/resource/Logarithm>”).

This lack of compliance with the Linked Data principles is not completely to blame on bad habits among the publishers of OpenMath CDs; it is also caused by technical and even conceptual problems:

- No MIME type for OpenMath objects/CDs has been specified. When publishing Linked Data, it is good practice to make both machine- and human-readable descriptions of the same things available from their URIs (cf. principle 3 and [6]). That is, from the URI of an OpenMath CD, *both* the CD XML file *and* its human-friendly HTML rendering should be available. We could even make an RDF description of the CD available, as RDF is most widely understood by Linked Data clients, and as we have machinery for translating OpenMath CDs to RDF (cf. [17]). The client indicates the desired data encoding by requesting a particular MIME type in the *Accept* header of its HTTP request; this mechanism is called *content negotiation* (cf. [23]). In

⁹ By the same argument as for principle 3, the DLMF content is merely *machine-readable*, but not *machine-understandable*, but, on the other hand, the ABRAMOWITZ/STEGUN book is even only *human-readable*.

accordance with best practices, I suggest *application/openmath+xml* to be introduced for the XML encoding of OpenMath.¹⁰

- CDs are not really meant as *machine*-understandable descriptions of symbols. They are mainly intended as somewhat rigorous descriptions for those *humans* who implement phrasebooks, i.e. translations between OpenMath Objects and the native languages of, e.g., computer algebra systems. This view is encouraged by the OpenMath 2.0 specification, which says “It is important to stress that it is not Content Dictionaries themselves which are being transmitted, but some ‘mathematics’ whose definitions are held within the Content Dictionaries.” [9] This view is plainly wrong on the Web of Data! By the “meta” CD, there is at least a well-meant approach to communicating CDs as OpenMath Objects¹¹. Most OpenMath-aware software, except a few editors, usually only supports OpenMath objects (= formulæ), but not CDs. For Linked Data applications, that has to change, and actually that change is not hard to make, because the CD XML format is well-specified and easy to implement.
- The semantics of FMPs is too weak. The application scenario outlined in section 2 assumes that FMPs carry definitions of symbols, but FMPs are not required to do so, as they could also carry asserted properties of symbols. Definitional FMPs have been discussed throughout the last 10 years (see, e.g., [10]) but still have not made it into the OpenMath standard. We might even need to go one step further¹² and introduce a notion of *computational* FMPs, as, for example, implicit definitions are not useful for term rewriting either. On the other hand, studying the practice of RDF-based Linked Data gives some hope, as, there, certain vocabulary terms are also used more liberally than they have been specified. The *rdfs:seeAlso* relation is semantically very weak (“used to indicate a resource that might provide additional information about the subject resource” [8]), but when used with Linked Data, it is commonly assumed that it points at a URI that is again machine-comprehensible and contains further Linked-Data-compliant information. Conversely, the *owl:sameAs* relation is commonly used to declare that two things, despite having different URIs, are the same (cf. [25]) – but hardly any Linked Data application makes use of the rest of the description logic based OWL ontology language, where this relation comes from. Similarly, there are practical applications of OpenMath CDs, such as STRATFORD’s and DAVENPORT’s unit converter [26], which simply *assume* that FMPs having the current symbol as the first argument of *relation1#eq* are definitional.
- There is no mechanism for linking OpenMath symbols to anything else but other OpenMath symbols (e.g. to DBpedia data). The latter is done via FMPs, but for the former we would have to be able to create typed links

¹⁰ It is subject to further discussion in the community whether the same MIME type should be used both for OpenMath objects and CDs.

¹¹ ... which even comes into operation in SCSCP [13]

¹² according to personal communication with MICHAEL KOHLHASE

from OpenMath symbols to arbitrary URIs. If the format restrictions for OpenMath symbol URIs were overcome (see next item), this *could* be done by encoding RDF links as FMPs, as we have proposed in [18]. Alternatively, one could permit RDFa metadata (a syntax for directly embedding RDF into XML-based languages) in CDs (see also [18]), but that would be a more intrusive change of the CD format.

- From a Linked Data point of view, the OpenMath schema of symbol URIs being constructed as `cdbase / cd # name` is too restrictive and should be liberalized. Not only are there “legacy” URIs having other formats out there on the Web of Data, but also every data publisher may have good reasons not to choose “hash URIs” (see, e. g., [5]). As processing everything after the `#` is up to the client, the consequent use of hash URIs for OpenMath symbols forces clients to always download a complete CD from the server, in which they would then have to locate the symbol with the desired name (e. g. using the `/CD/CDDefinition[Name = ...]` XPath expression). However, for a large CD, of which a client is only interested in few symbols, it would be more efficient to use “slash URIs”, such as <http://cdba.se/cd/name>, or similar formats, such as the MMT URIs under development for OMDoc [16]. In my opinion, it was also wrong to impose OpenMath’s strict `cdbase / cd # name` schema on Content MathML by de facto deprecating the liberal `csymbol/@definitionURL` attribute, which is now only permitted in non-strict markup [2].

I suggest that the OpenMath 3 specification address the conceptual issues and provide practical guidelines on how to address the technical issues.

4 Conclusion and Future Work

Looking at the state of the art of Linked Data and Linked Open datasets, we have identified a lack of mathematical semantics. We have pointed out how applications would benefit from mathematically annotated Linked Data and suggested OpenMath CDs, in combination with the prevalent RDF, to be used for that purpose. That, however, poses a number of technical and conceptual requirements on the OpenMath community, which I have described in detail, and which should be addressed in OpenMath 3.

As a particularly promising future research task, we have identified the integration of OpenMath-based computations right into queries against RDF-based Linked datasets. Consider querying a statistical dataset for the region with the highest increase of population density: Currently, that would require one step of querying (and obtaining population and area values), another step of computation (of the population densities), and a second step of querying (finding the maximum density). Or reconsider the example of computing a derived value in a dataset from section 2, where an expression is rewritten using the OpenMath definition of a function: When the arguments of that function are again derived values, we would also have to execute a chain of RDF queries and OpenMath-based

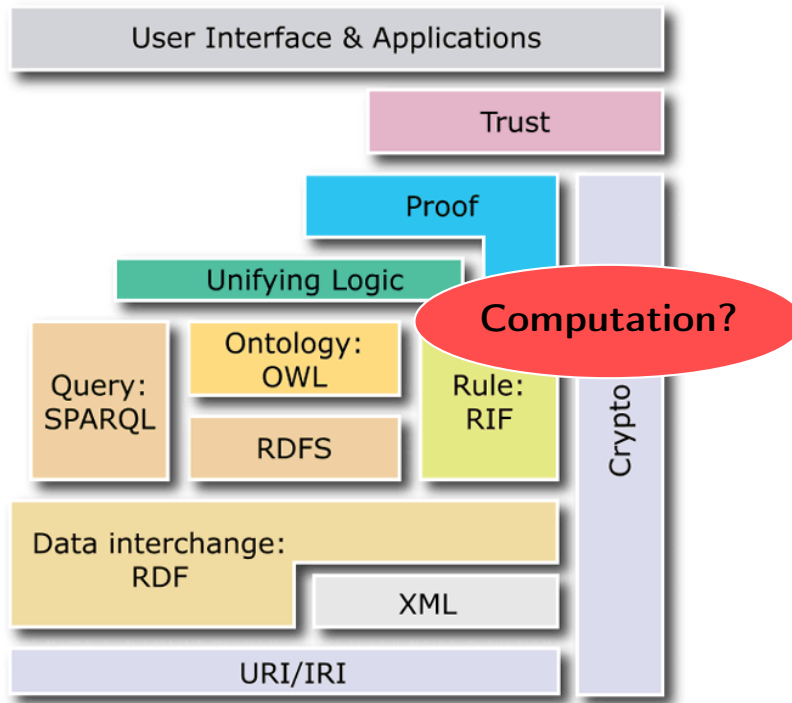


Fig. 2. The Semantic Web Layer Cake (originally by BERNERS-LEE)

term rewritings. RDF queries are usually made in the SPARQL language [22]. The SPARQL specification foresees the extension of the basic language by additional entailment regimes [14], which make a query return additional, *entailed* results, beyond the information that is explicitly encoded in the RDF graph being queried. The possibilities for an OpenMath entailment regime should be investigated. More pragmatically, and disregarding the consequences for computational complexity, many implementations of SPARQL query processors allow for defining extension functions, and some basic mathematical extension functions have already been implemented for certain query processors; it should be investigated how that can be generalized to arbitrary functions defined by OpenMath CDs. The goal to be pursued with that is to make *computation* an adequate part of the well-known Semantic Web layer cake (figure 2).

Acknowledgments: The original idea for integrating OpenMath CDs into the Web of Data and further steps, such as combining SPARQL and OpenMath, emerged from discussions with DENNY VRANDEČIĆ and a joint publication [28]. I would also like to thank DAVID CARLISLE, JAN WILLEM KNOPPER and MICHAEL KOHLHASE for their input.

References

- [1] M. Abramowitz and I. A. Stegun, eds. *Handbook of Mathematical Functions (with Formulas, Graphs, and Mathematical Tables)*. Applied Mathematics Series. National Bureau of Standards, 1964.
- [2] R. Ausbrooks, S. Buswell, D. Carlisle, G. Chavchanidze, S. Dalmas, S. Devitt, A. Diaz, S. Dooley, R. Hunter, P. Ion, M. Kohlhase, A. Lazrek, P. Libbrecht, B. Miller, R. Miner, M. Sargent, B. Smith, N. Soiffer, R. Sutor, and S. Watt. *Mathematical Markup Language (MathML) Version 3.0*. W3C Candidate Recommendation of 15 December 2009. World Wide Web Consortium, 2009. URL: <http://www.w3.org/TR/MathML3>.
- [3] S. Autexier, J. Campbell, J. Rubio, V. Sorge, M. Suzuki, and F. Wiedijk, eds. *Intelligent Computer Mathematics, 9th International Conference, AISC 2008 15th Symposium, Calculemus 2008 7th International Conference, MKM 2008 Birmingham, UK, July 28 - August 1, 2008, Proceedings*. LNAI 5144. Springer Verlag, 2008.
- [4] T. Berners-Lee. *Design Issues: Linked Data*. 2006. URL: <http://www.w3.org/DesignIssues/LinkedData.html> (visited on 01/20/2010).
- [5] D. Berrueta and J. Phipps. *Best Practice Recipes for Publishing RDF Vocabularies*. W3C Working Group Note. World Wide Web Consortium (W3C), Aug. 28, 2008. URL: <http://www.w3.org/TR/2008/NOTE-swbp-vocab-pub-20080828/>.
- [6] C. Bizer, R. Cyganiak, and T. Heath. *How to Publish Linked Data on the Web*. July 27, 2007. URL: <http://sites.wiwi.fu-berlin.de/suhl/bizer/pub/LinkedDataTutorial/20070727/>.
- [7] C. Bizer, T. Heath, and T. Berners-Lee. “Linked Data – The Story So Far”. In: *International Journal on Semantic Web and Information Systems* (2009). Special Issue on Linked Data.
- [8] D. Brickley and R. V. Guha. *RDF Vocabulary Description Language 1.0: RDF Schema*. W3C Recommendation. World Wide Web Consortium (W3C), Feb. 2004. URL: <http://www.w3.org/TR/2004/REC-rdf-schema-20040210/>.
- [9] S. Buswell, O. Caprotti, D. P. Carlisle, M. C. Dewar, M. Gaetano, and M. Kohlhase. *The Open Math Standard, Version 2.0*. Tech. rep. The Open Math Society, 2004. URL: <http://www.openmath.org/standard/om20>.
- [10] O. Caprotti and M. Oostdijk. “On Communicating Proofs in Interactive Mathematical Documents”. In: *Proceedings of Artificial Intelligence and Symbolic Computation, AISC’2000*. Ed. by E. R. Lozano. LNAI 1930. Springer Verlag, 2001, pp. 53–64.
- [11] R. Cyganiak and A. Jentzsch. *About the Linking Open Data dataset cloud*. July 14, 2009. URL: <http://richard.cyganiak.de/2007/10/lod/>.
- [12] S. Freundt, P. Horn, A. Kononov, S. Linton, and D. Roozmond. “Symbolic Computation Software Composability”. In: *Intelligent Computer Mathematics, 9th International Conference, AISC 2008 15th Symposium, Calculemus 2008 7th International Conference, MKM 2008 Birmingham*,

- UK, July 28 - August 1, 2008, *Proceedings*. Ed. by S. Autexier, J. Campbell, J. Rubio, V. Sorge, M. Suzuki, and F. Wiedijk. LNAI 5144. Springer Verlag, 2008, pp. 285–295.
- [13] S. Freundt, P. Horn, A. Kononov, S. Linton, and D. Roozemon. *Symbolic Computation Software Composability Protocol (SCSCP)*. Version 1.3. Mar. 27, 2009. URL: <http://www.symbolic-computation.org/SCSCP>.
 - [14] B. Glimm and B. Parsia. *SPARQL 1.1 entailment regimes*. W3C Working Draft. World Wide Web Consortium (W3C), June 1, 2010. URL: <http://www.w3.org/TR/2009/WD-sparql11-entailment-20100601/>.
 - [15] O. Hartig, H. Mühleisen, and J.-C. Freytag. “Linked Data for Building a Map of Researchers”. In: *Scripting and Development for the Semantic Web (SFSW2009)*. Ed. by C. Bizer, S. Auer, and G. A. Grimnes. CEUR Workshop Proceedings 449. May 2009. URL: <http://CEUR-WS.org/Vol-449/>.
 - [16] M. Kohlhase, F. Rabe, and V. Zholudev. “Towards MKM in the Large: Modular Representation and Scalable Software Architecture”. In: *Intelligent Computer Mathematics*. LNAI 6167. in press. Springer Verlag, 2010. URL: <http://kwarc.info/kohlhase/papers/mkm10-scalable.pdf>.
 - [17] C. Lange. “wiki.openmath.org – how it works, how you can participate”. In: *22nd OpenMath Workshop*. Ed. by J. H. Davenport. July 2009. arXiv: 1003.5192.
 - [18] C. Lange and M. Kohlhase. “A Mathematical Approach to Ontology Authoring and Documentation”. In: *MKM/Calculemus 2009 Proceedings*. Ed. by J. Carette, L. Dixon, C. Sacerdoti Coen, and S. M. Watt. LNAI 5625. Springer Verlag, July 2009, pp. 389–404. URL: <https://svn.omdoc.org/repos/omdoc/trunk/doc/blue/foaf/mkm09.pdf>.
 - [19] M. Marchiori. “The Mathematical Semantic Web”. In: *Mathematical Knowledge Management, MKM’03*. Ed. by A. Asperti, B. Buchberger, and J. H. Davenport. LNCS 2594. Keynote. Springer Verlag, 2003.
 - [20] National Institute of Standards and Technology, ed. *Digital Library of Mathematical Functions*. May 7, 2010. URL: <http://dlmf.nist.gov>.
 - [21] T. Omitola, C. L. Koumenides, I. O. Popov, Y. Yang, M. Salvadores, M. Szomszor, T. Berners-Lee, N. Gibbins, W. Hall, mc schraefel, and N. Shadbolt. “Put in Your Postcode, Out Comes the Data: A Case Study”. In: *ESWC*. Ed. by L. Aroyo, G. Antoniou, E. Hyvönen, A. ten Teije, H. Stuckenschmidt, L. Cabral, and T. Tudorache. Lecture Notes in Computer Science 6088. Springer, June 2010, pp. 318–332.
 - [22] E. Prud’hommeaux and A. Seaborne. *SPARQL Query Language for RDF*. W3C Recommendation. World Wide Web Consortium (W3C), Jan. 2008. URL: <http://www.w3.org/TR/2008/REC-rdf-sparql-query-20080115/>.

- [23] L. Sauermann and R. Cyganiak. *Cool URIs for the Semantic Web*. W3C Working Draft. World Wide Web Consortium (W3C), Dec. 3, 2008. URL: <http://www.w3.org/TR/2008/WD-cooluris-20081203/>.
- [24] F.-P. Servant. “Linking Enterprise Data”. In: *Linked Data on the Web (LDOW 2008)*. Ed. by C. Bizer, T. Heath, K. Idehen, and T. Berners-Lee. CEUR Workshop Proceedings 369. Apr. 2008. URL: <http://CEUR-WS.org/Vol-369/>.
- [25] M. K. Smith, C. Welty, and D. L. McGuinness. *OWL Web Ontology Language Guide*. W3C Recommendation. World Wide Web Consortium (W3C), Feb. 2004. URL: <http://www.w3.org/TR/2004/REC-owl-guide-20040210/>.
- [26] J. Stratford and J. H. Davenport. “Unit Knowledge Management”. In: *Intelligent Computer Mathematics, 9th International Conference, AISC 2008 15th Symposium, Calculemus 2008 7th International Conference, MKM 2008 Birmingham, UK, July 28 - August 1, 2008, Proceedings*. Ed. by S. Autexier, J. Campbell, J. Rubio, V. Sorge, M. Suzuki, and F. Wiedijk. LNAI 5144. Springer Verlag, 2008, pp. 382–397.
- [27] D. Vrandečić, M. Krötzsch, S. Rudolph, and U. Lösch. “Leveraging Non-Lexical Knowledge for the Linked Open Data Web”. In: *The Fifth RAFT’2010, the yearly bilingual publication on nonchalant research*. Ed. by R. Héliot and A. Zimmermann. URL: http://km.aifb.kit.edu/projects/numbers/linked_open_numbers.pdf.
- [28] D. Vrandečić, C. Lange, M. Hausenblas, J. Bao, and L. Ding. “Semantics of Governmental Statistics Data”. In: *Proceedings of WebSci’10: Extending the Frontiers of Society On-Line*. 2010. URL: <http://journal.webscience.org/400/>.
- [29] Wikimedia Foundation, ed. *Linked Data (from Wikipedia, the free encyclopedia)*. June 18, 2010. URL: http://en.wikipedia.org/w/index.php?title=Linked_Data&oldid=368472510.